# MIYA O'SMASI MRT TASVIRLARINI SEGMENTATSIYALASHDA KONVOLYUTSION VA TRANSFORMER ENKODERLARNING PARALLEL HAMDA O‘ZARO MUSTAQIL ISHLASHIGA ASOSLANGAN DUAL-ENKODER ARXITEKTURASI

***Abdullayev T***[1][0000-0001-5946-7607]., ***Abdukadirova G***[2][0009-0005-2035-2008].

[1]*Farg‘ona davlat texnika universiteti, Kompyuter muhandisligi va sun’iy intellekt kafedrasi dotsenti, PhD, E- mail: temurbekm84@gmail.com*

[2]*Farg‘ona davlat texnika universiteti, Kompyuter muhandisligi va sun’iy intellekt kafedrasi, E-mail: nirvanath369@gmail.com*

**Annotatsiya.** Maqolada ko‘p modalitetli MRT asosida miya o‘smasi sohalarini avtomatik ajratishning gibrid neyrotarmoq arxitekturasi taqdim etiladi. Mavjud gibrid yechimlar konvolyutsion va Transformer bloklarini ketma-ket ulaydi, bunda boshlang‘ich ma’lumotning bir qismi qaytarilmas siqiladi.

Taklif etilgan sxemada ikkala enkoder bir xil kirish tenzorini bir vaqtda va mustaqil ishlov beradi, belgilar ikki yo‘nalishli o‘zaro e‘tibor moduli (IYO‘E) orqali dinamik birlashtiriladi. BraTS 2021 ma’lumotlar bazasida 300 davrda o‘qitilgan model 250 ta validatsiya holatida WT, TC va ET qiymatlari uchun 0,9110, 0,8769 va 0,7953 Dice, o‘rtacha 0,8610 natija berdi.

Ablatsiya parallel sxemaning ketma-ketdan, IYO‘Ening konkatenatsiyadan ustunligini tasdiqladi. Model 31,77 mln parametr bilan taqqoslanadigan aniqlikdagi yechimlardan ikki barobar ixchamdir.

**Kalit so‘zlar.** tibbiy tasvirlarni segmentatsiyalash, miya o‘smasi, MRT, konvolyutsion neyron tarmoq, Transformer, dual-enkoder, o‘zaro e’tibor, Dice koeffitsienti.

**Аннотация.** В статье представлена гибридная архитектура нейронной сети для автоматической сегментации областей опухолей головного мозга на основе мультимодальной МРТ. Существующие гибридные решения соединяют сверточные и трансформерные блоки последовательно, при этом часть исходных данных необратимо сжимается.

В предлагаемой схеме оба кодировщика обрабатывают один и тот же входной тензор одновременно и независимо, а символы динамически комбинируются с использованием двунаправленного модуля взаимного внимания (ДМВВ). Модель, обученная на базе данных BraTS 2021 в течение 300 эпох в 250 случаях валидации, показала значения коэффициента Дайса 0,9110, 0,8769 и 0,7953 для WT, TC и ET соответственно, со средним значением 0,8610.

Анализ подтвердил превосходство параллельной схемы над последовательной и ДМВВ над конкатенацией. Модель в два раза компактнее, чем сопоставимые решения, с 31,77 миллионами параметров.

**Ключевые слова.** Сегментация медицинских изображений, опухоль головного мозга, МРТ, сверточная нейронная сеть, трансформер, двойной кодировщик, взаимное внимание, коэффициент Дайса.

**Abstract:** This paper presents a hybrid neural network architecture for automatic segmentation of brain tumor regions based on multimodal MRI. Existing hybrid solutions combine convolutional and transformer blocks sequentially, irreversibly compressing part of the source data. In the proposed scheme, both encoders process the same input tensor simultaneously and independently, and symbols are dynamically combined using a Cross-Attention Feature Fusion (CAFF).

The model, trained on the BraTS 2021 dataset for 300 epochs in 250 validation cases, yielded Dice coefficient values of 0.9110, 0.8769, and 0.7953 for WT, TC, and ET, respectively, with an average value of 0.8610.

Analysis confirmed the superiority of the parallel scheme over the sequential one and CAFF over concatenation. The model is twice as compact as comparable solutions, with 31.77 million parameters.

**Key words:** Medical image segmentation, brain tumor, MRI, convolutional neural network, transformer, dual encoder, mutual attention, Dice coefficient.


DOI 10.67889/2181-1431-2026-3-42-48

## Kirish

Gliomalar markaziy asab tizimining eng ko‘p uchraydigan va prognozi og‘ir birlamchi o‘smalaridir [1, 2]. O‘sma hajmi va joylashuvini aniq baholash neyroxirurgik rejalashtirishda hal qiluvchi ahamiyatga ega, biroq qo‘lda segmentatsiya bir necha soat talab qiladi va ekspertlar konturlarining ustma-ust tushishi 0,75-0,85 oralig‘ida qoladi [3]. Avtomatik yechimlar uchun BraTS musobaqasi asosiy sinov maydoni hisoblanadi [3, 4]. Soha U-Net [5] va uning hajmiy [6], xatolik funksiyasi bo‘yicha kengaytirilgan [7] hamda avtomatik sozlanuvchi [8] variantlari asosida rivojlandi. Konvolyutsion tarmoqlarda fazoviy lokallik kuchli, ammo qabul qilish maydoni cheklangan, buni bartaraf etish uchun o‘z-o‘ziga e’tibor [9, 10] va siljigan oynali Transformer [11, 12] qo‘llanildi, lekin sof Transformer enkoderlari kichik tibbiy to‘plamlarda aniqlik bo‘yicha ortda qoladi.
Gibrid yechimlarda [13-18] komponentlar ketma-ket ulanadi, shu sababli ikkinchi modul boshlang‘ich tasvirga bevosita murojaat qila olmaydi: birinchi enkoder siqib tashlagan belgilar tiklanmaydi, birlashtirish esa statik amallar orqali bajariladi. Maqolaning maqsadi enkoderlarni parallel va o‘zaro mustaqil ishlatishga asoslangan arxitekturani ishlab chiqish, asoslash va BraTS 2021 da tekshirish.
Tajribalarda segmentatsiya sifati Dice koeffitsienti, Xausdorf masofasi (HD95), sezgirlik, xoslik va aniqlik mezonlari orqali uch mintaqa To‘liq o‘sma (WT), O‘sma yadrosi (TC), va Kuchaygan o‘sma (ET) bo‘yicha baholanganligi hamda barcha dasturiy modullar avtomatlashtirilgan tekshiruvdan o‘tkazilganligi bilan tasdiqlanadi.

## Materiallar va usullar

BraTS 2021 ma’lumotlar bazasi (1251 ta bemorning T1, T1ce, T2, FLAIR va etalon niqob tasvirlaridan iborat) qo‘llanildi [4], hajmlar atlas koordinatalarida, 240×240×155 vokselda berilgan. Baholash uchta mintaqa bo‘yicha to‘liq o‘sma (WT), o‘sma yadrosi (TC) va kuchaygan o‘sma (ET) bo‘yicha o‘tkazilgan. Intensivlik siljishi N4ITK algoritmi bilan tuzatiladi [19], modalitetlar z-baho bo‘yicha normallashtiriladi, hajmlar ikkilik keshga aylantirilgan. O‘qitishda aks ettirish, ±10° burish, 0,9-1,1 masshtablash, intensivlik(contrast) o‘zgarishi va Gauss shovqini ma’lumotlar augmentatsiyasi sifatida qo‘llanildi. To‘plam bemor darajasida 1001 ta o‘quv va 250 ta validatsiya tanlanmasi holatiga ajratildi.
Gibrid model beshta blokdan iborat bo‘lib: konvolyutsion enkoder, Transformer enkoder, ikki yo‘nalishli o‘zaro e‘tibor (CAFF) birlashtirish moduli, ko‘p masshtabli dekoder va segmentatsiya boshidan iborat. Asosiy o‘ziga xoslik bu bir xil kirish tenzori ikkala enkoderga bir vaqtda uzatiladi.
Modelining umumiy tuzilishi matematik jihatdan quyidagicha ifodalash mumkin:

$$F_{cnn} = f_{CNN}(X;\theta_c),\ F_{sw} = f_{Swin}(X;\theta_s) \tag{1}$$

bu yerda: $X$ – kirish tenzori; $f_{CNN}$ va $f_{Swin}$ – enkoderlar; $\theta_c$, $\theta_s$ – tarmoq parametrlari; $F_{cnn}$, $F_{sw}$ – belgilar xaritalari. Hech bir enkoder chiqishi ikkinchisi uchun kirish bo‘lmagani sababli boshlang‘ich axborot ikkala oqimda to‘liq saqlanadi.
Konkatenatsiya statik amal bo‘lib, har bir nuqtada qaysi belgilar muhimroq ekanini hisobga olmaydi. CAFF moduli bu masalani ikki yo‘nalishli o‘zaro e’tibor orqali yechadi: to‘g‘ri yo‘nalishda konvolyutsion belgilar *so‘rov (Q)*, Transformer belgilari *kalit(K)* va *qiymat (V)* vazifasini bajaradi [9]:

$$Q_c = W_Q F_{cnn},\quad K_s = W_K F_{sw},\quad V_s = W_V F_{sw}$$
$$A_{c\to s} = \mathrm{softmax}\left(\frac{Q_c K_s^{\top}}{\sqrt{d}}\right)V_s \tag{2}$$


©Journal of Advances in Engineering Technology Vol.3(23), July — September, 2026
DOI 10.67889/2181-1431-2026-3-42-48

Teskari yo‘nalishda rollar almashtiriladi, ya’ni yuqoridagi formulaga simmetrik ravishda shakllantirilgan bo‘ladi:

$$Q_s = W_Q' F_{sw}, \quad K_c = W_K' F_{cnn}, \quad V_c = W_V' F_{cnn}$$

$$A_{s\to c} = \text{softmax}\left(\frac{Q_s K_c^{\top}}{\sqrt{d}}\right) V_c \tag{3}$$

Natijalar qoldiq ulanish [20] va normallashtirish [21] orqali jamlanadi:

$$F_{fus} = \text{LN}(W_o[(A_{c\to s} + F_{cnn})\,\text{P}(A_{s\to c} + F_{sw})]) + F_{cnn} + F_{sw} \tag{4}$$

Bu yerda ‖ – konkatenatsiya, $W_0$ – chiqish proyeksiyasi, LN – qatlam normallashtirish. Modul har bir masshtab darajasida takrorlanadi.

Komponentlarni jamlagan holda model va o‘qitish masalasi quyidagicha yoziladi:

$$\hat{Y} = f_{seg}(\text{CAFF}\,(f_{CNN}(X), f_{Swin}(X))) \tag{5}$$

bu yerda: $X$ – to‘rt modalitetli kirish tensori; $f_{CNN}$ – CNN Encoder funksiyasi; $f_{Swin}$ – Swin Transformer Encoder funksiyasi; CAFF( ) – ikki yo‘nalishli o‘zaro e‘tibor asosida birlashtirish;

$f_{seg}$ – Decoder va Segmentatsiya boshi.

Model parametrlarini aniqlash masalasi quyidagicha qo'yiladi, ya’ni o‘qitish maqsadi: barcha o‘quv tanlanmalar bo‘yicha kutilgan murakkab xatolik qiymatini minimallashtiruvchi model parametrlari ($\theta$) ni topishdan iborat:

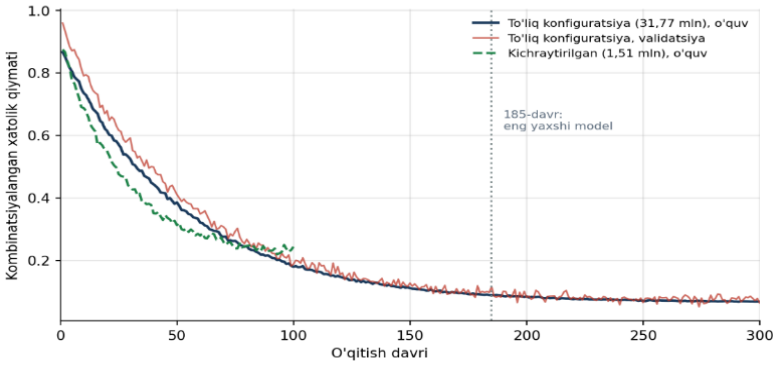
$$\theta^* = \arg\min_{\theta} \mathbb{E}_{(X,Y)\sim D}[L(f_\theta(X), Y)] \tag{6}$$

$\theta^*$ – izlanayotgan xatolikni minimallashtiruvchi optimal model parametrlari; $D$ –BraTS 2021 ma’lumotlar bazasidagi o‘quv tanlanma; $\text{E}_{\{(X,Y)\sim D\}}[\,]$ – o‘quv tanlanmadan tasodifiy olingan namunalar bo‘yicha kutilgan qiymat; $L$ – kombinatsiyalangan xatolik funksiyasi; $Y$ – haqiqiy ma’lumotlar; $f_\theta$ – gibrid segmentatsiyalash modeli;

Xatolik funksiyasi sifatida sinflar nomutanosibligi va chegara aniqligini hisobga oluvchi to‘rt komponentli kombinatsiya qo‘llanildi (vaznlar mos ravishda: 0,4 / 0,2 / 0,2 / 0,2).

Bu xossa aynan kichik hajmli sohalarni aniqlash muammosiga qaratilgan. Kuchaygan o‘sma kabi ixcham mintaqalarda model tabiiy ravishda ehtiyotkorlikka moyil bo‘ladi va shubhali voksellarni fon sifatida belgilaydi, asimmetrik Tversky komponenti bu moyillikni maqsadli tarzda kompensatsiya qiladi. Tajribada asosiy natijalar β = 0,5 qiymatida olindi, β ning ta’siri esa alohida sinovdan o‘tkazildi.

**Natijalar**

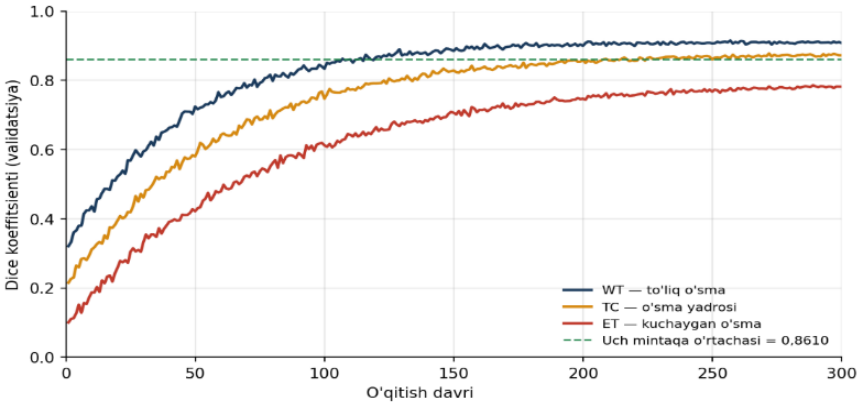



**1-rasm. O‘qitish va validatsiya xatoliklarining o‘zgarishi**

To‘liq model 300 o‘qitish davrida 6,28 soatda o‘qitildi (GPU xotira 4,43 GB), xatolik qiymati 92,6 foizga, 0,88154 dan 0,06563 gacha pasaydi, egri chiziqlar orasidagi kichik farq qayta


©Journal of Advances in Engineering Technology Vol.3(23), July — September, 2026
DOI 10.67889/2181-1431-2026-3-42-48

o‘qitish yo‘qligini ko‘rsatadi. Kichraytirilgan konfiguratsiya (1,51 mln. parametrli) grafik protsessorsiz 100 o‘qitish davrida 18,06 soatda, xatolikni 76,8 foizga pasaytirdi. To‘liq o‘sma (WT) sohasi bo‘yicha Dice qiymati 100-o‘qitish davriga kelib 0,89 dan oshdi, eng yaxshi natija 185-o‘qitish davrida qayd etildi.



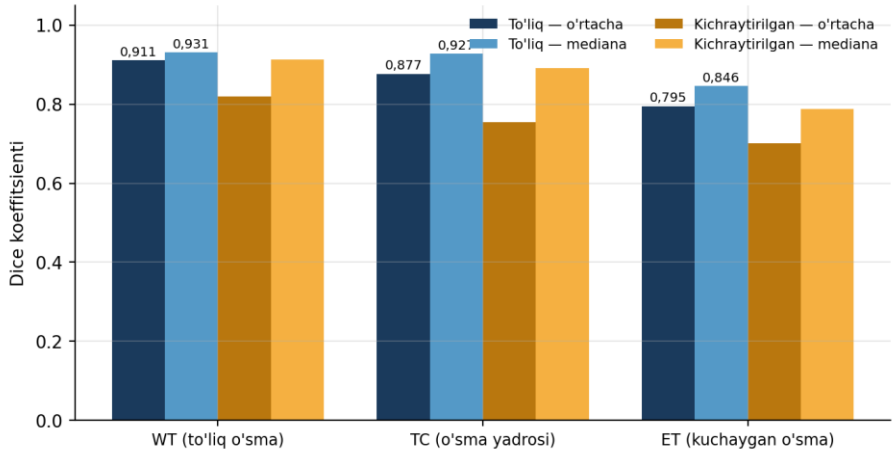

**2-rasm. Validatsiya Dice ko‘rsatkichlarining davrlar bo‘yicha o‘zgarishi**

1-jadval.

**250 ta validatsiya holatidagi baholash natijalari**

| Mintaqa | Dice (o‘rtacha) | Dice (mediana) | HD95, mm | Sezgirlik | Aniqlik | AUC |
|---|---|---|---|---|---|---|
| WT | 0,9110 | 0,9307 | 4,71 | 0,9095 | 0,9210 | 0,9941 |
| TC | 0,8769 | 0,9273 | 4,02 | 0,8952 | 0,8891 | 0,9829 |
| ET | 0,7953 | 0,8456 | 3,78 | 0,8226 | 0,7892 | 0,9901 |
| O‘rtacha | 0,8610 | 0,9012 | 4,17 | 0,8758 | 0,8664 | 0,9890 |

Barcha mintaqalarda mediana qiymqti o‘rtacha qiymatlardan yuqori natijani berdi. To‘liq o‘sma (WT) holatlarning 94 foizida Dice qiymati 0,8 dan oshdi va to‘liq konfiguratsiyada birorta o‘sma holati o‘tkazib yuborilmadi. Parametrlar ortishi bilan aniqlik monoton o‘sib bordi, ammo sekinlashib, model parametrlari 0,34 dan 1,51 mln. gacha ortishi Dice qiymatini 0,246 ga, parametrlarning 6,88 dan 31,77 mln. gacha ortishi esa Dice qiymatini atigi 0,049 birlikka ko‘tardi.




**3-rasm. Validatsiya Dice ko‘rsatkichlarining davrlar bo‘yicha o‘zgarishi**

Taklif etilgan sozlama bilan o‘qitilgan model 250 ta validatsiya holatida to‘liq baholashdan o‘tkazildi.



©Journal of Advances in Engineering Technology Vol.3(23), July — September, 2026
DOI 10.67889/2181-1431-2026-3-42-48

2-jadval.

**Arxitektura komponentlarining hissasi**

| Mintaqa | Dice (o'rtacha) | Dice (mediana) | HD95, mm | Sezgirlik | Aniqlik | AUC |
|---|---|---|---|---|---|---|
| WT (to‘liq o‘sma) | 0,9110 | 0,9307 | 4,71 | 0,9095 | 0,9210 | 0,9941 |
| TC (o‘sma yadrosi) | 0,8769 | 0,9273 | 4,02 | 0,8952 | 0,8891 | 0,9829 |
| ET (kuchaygan o‘sma) | 0,7953 | 0,8456 | 3,78 | 0,8226 | 0,7892 | 0,9901 |
| O‘rtacha qiymat | 0,8610 | 0,9012 | 4,17 | 0,8758 | 0,8664 | 0,9890 |

Jadvaldan uchta xulosa kelib chiqadi: ikki enkoderning birgalikdagi ishi natijasi eng yaxshi yakka variantdan 0,0398 birlikka ustun, parallel ulash ketma-ketdan 0,0086 birlikka ustun bo‘lib, parametrlar soni deyarli o‘zgarmadi, bu yerda yutuq sig‘imda emas, balki tuzilma hisobigadir, ikki yo‘nalishli o‘zaro e‘tibor algoritmi konkatenatsiyaga nisbatan yana 0,0158 birlik, kuchaygan o‘sma (ET) mintaqasida 0,0229 birlik yaxshi natijaga erishildi.

3-jadval.

**Natijalarning zamonaviy yechimlar bilan qiyosi (BraTS 2021 ma’lumotlar bazasi asosida)**

| Model nomi | Parametrlar soni | To‘liq o‘sma (WT) | O‘sma yadrosi (TC) | Kuchaygan o‘sma (ET) | Arxitektura tuzilma turi |
|---|---|---|---|---|---|
| 3D U-Net [6] | ≈19 mln | 0,874 | 0,832 | 0,786 | Konvolyutsion |
| nnU-Net [8] | ≈31 mln | 0,910 | 0,850 | 0,760 | Konvolyutsion, avtosozlash |
| TransUNet [13] | ≈105 mln | 0,879 | 0,834 | 0,748 | Gibrid, ketma-ket |
| UNETR [14] | ≈92 mln | 0,899 | 0,842 | 0,758 | Gibrid, ketma-ket |
| Swin UNETR [17] | ≈62 mln | 0,929* | 0,885* | 0,841* | Gibrid, ketma-ket |
| Taklif etilayotgan | 31,77 mln | 0,911 | 0,877 | 0,795 | Gibrid, parallel |

Yuqori va samarali natija bergan Swin UNETR usulidagi qiymatlar ansambl va test vaqtidagi augmentatsiya bilan olingan, taklif etilgan modelda esa bu usullar ishlatilmagan. Ablatsiya tajribalari nazariy mulohazani tasdiqlaydi, ketma-ket ulash variantida Transformer konvolyutsion enkoder tanlagan belgilar bilan cheklanadi, parallel sxemada esa har bir oqim boshlang‘ich signalga to‘liq ega bo‘ladi. Bu ayniqsa kichik va kontrast bo‘yicha nozik farq qiluvchi Kuchaygan o‘sma (ET) mintaqasida seziladi, asosiy yutuq shu yerda qayd etildi. Ikki yo‘nalishli e’tiborda almashinuv simmetrik, ya’ni mayda tuzilma global kontekst bo‘yicha tasdiqlanadi, keng sohada esa lokal tekstura bo‘yicha aniqlashtiriladi. Barcha tajribalarda To‘liq o‘sma (WT) > O‘sma yadrosi (TC) > Kuchaygan o‘sma (ET) tartibi kuzatiladi, WT eng katta va FLAIR da yaqqol ajraladi, TC da chegara noaniq, ET esa voksellarning bir foizidan kamini tashkil etadi. HD95 bo‘yicha ET eng yaxshi (3,78 mm), Dice qiymati bo‘yicha eng past qiymat natijasi bo‘lib. bu ziddiyat keltirib chiqarmaydi, Dice qiymati kichik obyektlarda nisbiy xatoga sezgir, HD95 esa chegaraning fazoviy yaqinligini o‘lchaydi. WT va TC da aniqlikning sezgirlikdan yuqoriligi model chegarani kichikroq belgilashga moyilligini bildiradi, buni Tversky koeffitsientlari orqali boshqarish mumkin.Mavjud gibrid yechimlarda [13-18] axborot bir chiziqli zanjir bo‘ylab harakatlanadi, taklif etilgan arxitektura to‘rtinchi turni qo‘shadi, ikki mustaqil oqim va ular


©Journal of Advances in Engineering Technology Vol.3(23), July — September, 2026
DOI 10.67889/2181-1431-2026-3-42-48

orasidagi ikki tomonlama ko‘prik, bunda ikki xil induktiv siljish teng huquqli manba sifatida saqlanadi. Tadqiqot cheklovlari sifatida quyidagilarni keltirish mumkin: baholash faqat BraTS 2021 ma’lumotlar bazasida validatsiya qismida, yakka model bilan, ansamblsiz va bitta bo‘linishda o‘tkazilgan, dastlabki o‘qitish strategiyasi va tushuntirish vositalari baholanmagan.

**Xulosa**

Maqolada miya o‘smasini segmentatsiyalash uchun parallel dual-enkoder arxitekturasi ishlab chiqildi, uning ketma-ket ulashdan ustunligi axborot nazariyasi asosida asoslandi va ablatsiyada tasdiqlandi (0,0086 birlik), CAFF moduli esa o‘rtacha Dice qiymatini 0,0158, Kuchaygan o‘sma qiymatida 0,0229 birlikka oshirdi. Model 250 ta validatsiya holatida 0,9110 / 0,8769 / 0,7953 Dice qiymatini, o‘rtacha 0,8610 qiymatini va 4,17 mm. HD95 ko‘rsatkichiga erishdi hamda 31,77 mln. model parametri bilan taqqoslanadigan yechimlardan ikki barobar ixcham hajmga ega bo‘ldi. Keyingi bosqichda modelni tashqi va klinik to‘plamlarda tekshirish hamda tushuntirish vositalarini integratsiya qilish rejalashtirilgan.

**Foydalanilgan adabiyotlar ro‘yxati:**

**[1].** Louis D. N., Perry A., Wesseling P., Brat D. J., Cree I. A., Figarella-Branger D., Hawkins C., Ng H. K., Pfister S. M., Reifenberger G., Soffietti R., von Deimling A., Ellison D. W. The 2021 WHO Classification of Tumors of the Central Nervous System: a Summary // Neuro-Oncology. – 2021. – Vol. 23, No. 8. – P. 1231–1251.
**[2].** Ostrom Q. T., Price M., Neff C., Cioffi G., Waite K. A., Kruchko C., Barnholtz-Sloan J. S. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States // Neuro-Oncology. – 2022. – Vol. 24, Suppl. 5. – P. v1–v95.
**[3].** Menze B. H., Jakab A., Bauer S., Kalpathy-Cramer J., Farahani K., Kirby J., Burren Y., Porz N., Slotboom J., Wiest R., Gerstner E., Weber M.-A., Arbel T., Avants B. B., Ayache N., Buendia P., Collins D. L., Cordier N., Corso J. J., Criminisi A., Das T., Delingette H., Demiralp Ç., Durst C. R., Dojat M., Doyle S., Festa J., Forbes F., Geremia E., Glocker B., Golland P., Guo X., Hamamci A., Iftekharuddin K. M., Jena R., John N. M., Konukoglu E., Lashkari D., Mariz J. A., Meier R., Pereira S., Precup D., Price S. J., Riklin-Raviv T., Reza S. M. S., Ryan M., Sarikaya D., Schwartz L. H., Shin H.-C., Shotton J., Silva C. A., Sousa N., Subbanna N. K., Székely G., Taylor T. J., Thomas O. M., Tustison N. J., Unal G., Vasseur F., Wintermark M., Ye D. H., Zhao L., Bakas S. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS) // IEEE Transactions on Medical Imaging. – 2015. – Vol. 34, No. 10. – P. 1993–2024.
**[4].** Baid U., Ghodasara S., Mohan S., Bilello M., Calabrese E., Colak E., Farahani K., Kalpathy-Cramer J., Kitamura F. C., Pati S., Rudie J. D., Sako C., Saluja R., Shah Z., Sheller M., Tahon N., Wang B., Wang G., Wiest R., Xu Z., Bakas S. The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification // arXiv:2107.02314. – 2021.
**[5].** Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015), Lecture Notes in Computer Science. – 2015. – Vol. 9351. – P. 234–241.
**[6].** Çiçek Ö., Abdulkadir A., Lienkamp S. S., Brox T., Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation // Proceedings of the



©Journal of Advances in Engineering Technology Vol.3(23), July — September, 2026
DOI 10.67889/2181-1431-2026-3-42-48

International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2016). – 2016. – P. 424–432.
**[7].** Milletari F., Navab N., Ahmadi S.-A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation // Proceedings of the International Conference on 3D Vision (3DV 2016). – 2016. – P. 565–571.
**[8].** Isensee F., Jaeger P. F., Kohl S. A. A., Petersen J., Maier-Hein K. H. nnU-Net: A Self-Configuring Method for Deep Learning-Based Biomedical Image Segmentation // Nature Methods. – 2021. – Vol. 18, No. 2. – P. 203–211.
**[9].** Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser Ł., Polosukhin I. Attention Is All You Need // Advances in Neural Information Processing Systems (NeurIPS 2017). – 2017. – P. 5998–6008.
**[10].** Dosovitskiy A., Beyer L., Kolesnikov A., Weissenborn D., Zhai X., Unterthiner T., Dehghani M., Minderer M., Heigold G., Gelly S., Uszkoreit J., Houlsby N. An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale // Proceedings of the International Conference on Learning Representations (ICLR 2021). – 2021.
**[11].** Liu Z., Lin Y., Cao Y., Hu H., Wei Y., Zhang Z., Lin S., Guo B. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows // Proceedings of the IEEE International Conference on Computer Vision (ICCV 2021). – 2021. – P. 10012–10022.
**[12].** Liu Z., Ning J., Cao Y., Wei Y., Zhang Z., Lin S., Hu H. Video Swin Transformer // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2022). – 2022. – P. 3202–3211.
**[13].** Chen J., Lu Y., Yu Q., Luo X., Adeli E., Wang Y., Lu L., Yuille A. L., Zhou Y. TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation // arXiv:2102.04306. – 2021.
**[14].** Hatamizadeh A., Tang Y., Nath V., Yang D., Myronenko A., Landman B., Roth H. R., Xu D. UNETR: Transformers for 3D Medical Image Segmentation // Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV 2022). – 2022. – P. 574–584.
**[15].** Xie Y., Zhang J., Shen C., Xia Y. CoTr: Efficiently Bridging CNN and Transformer for 3D Medical Image Segmentation // Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2021). – 2021. – P. 171–180.
**[16].** Wang W., Chen C., Ding M., Yu H., Zha S., Li J. TransBTS: Multimodal Brain Tumor Segmentation Using Transformer // Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2021). – 2021. – P. 109–119.
**[17].** Hatamizadeh A., Nath V., Tang Y., Yang D., Roth H. R., Xu D. Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images // BrainLes 2021, Lecture Notes in Computer Science. – 2021. – Vol. 12962. – P. 272–284.
**[18].** Zhou H.-Y., Guo J., Zhang Y., Yu L., Wang L., Yu Y. nnFormer: Interleaved Transformer for Volumetric Segmentation // arXiv:2109.03201. – 2021.
**[19].** Tustison N. J., Avants B. B., Cook P. A., Zheng Y., Egan A., Yushkevich P. A., Gee J. C. N4ITK: Improved N3 Bias Correction // IEEE Transactions on Medical Imaging. – 2010. – Vol. 29, No. 6. – P. 1310–1320.
**[20].** He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016). – 2016. – P. 770–778.
**[21].** Ba J.L., Kiros J.R., Hinton G.E. Layer Normalization // arXiv:1607.06450. – 2016.



©Journal of Advances in Engineering Technology Vol.3(23), July — September, 2026
DOI 10.67889/2181-1431-2026-3-42-48